A Review Paper: Speech Enhancement Using Different Spectral Subtraction Algorithms

Nisha Singh
M.Tech Student
Amity University
Noida
chauhan.nisha.in@gmail.com

Divya Gupta
Assistant Professor
Amity University
Noida
fromdivya@gmail.com

***Abstract* -This is a review paper that aims to provide an overview of Spectral Subtraction Algorithms that have been proposed for enhancement of speech degraded by additive background noise. Three major reasons of speech degradation are noise from background, resounding and speech from the attacking speakers. Spectral subtraction has undergone many changes. In this paper few modified spectral subtraction algorithms are briefly studied and reviewed. This paper briefly presents the comparison between some of spectral subtraction algorithms.**

Keywords -Speech Enhancement, Spectral Subtraction, Linear Spectral Subtraction, Comparison.

I. INTRODUCTION

Speech enhancement aims to remove or reduce background noise, Improve signal-to-noise ratio (SNR), Assumes stationary noise or at least that noise is more stationary than speech and a tradeoff between speech distortion and noise distortion (residual noise). In speech communication, the speech signal is always have some noise. In most cases the main component of noise is the background noise of the environment where the source of speech lies, that adds to the speech signal. The effect of this added noise is to make the listening task difficult for a listener who is listening directly, there are many more far reaching negative effects when we process the degraded speech for some other applications. A problem that is related is processing degraded speech in preparation for coding by a bandwidth compression system. Hence speech enhancement not only involves processing speech signals for human listening but also for further processing prior to listening. Speech enhancement mainly aims to improve the perceptual aspects of speech including overall quality, intelligibility, or degree of listener fatigue[4]. Research on speech enhancement techniques started more than 40 years ago at AT&T Bell Laboratories by Schroeder as mentioned in [9].Schroeder proposed an analog implementation of the spectral magnitude subtraction method. However, more than 15 later, the spectral subtraction method as proposed by Boll is a popular speech enhancement technique through noise reduction due to its simple underlying concept and its effectiveness in enhancing speech degraded by additive noise. The technique is based on the direct estimation of the short-term spectral magnitude[17]. Noise reduction or speech enhancement algorithms in general, attempt to improve the performance of communication systems when their input or output signals are corrupted by noise [17]. The main objective of speech enhancement is to improve one or more perceptual aspects of speech, such as the speech quality or intelligibility. In this paper speech enhancement methods are reviewed in detail.[4]

II. SPECTRAL SUBTRACTION METHODS

There are various spectral subtraction methods or algorithms proposed for improving the noise quality and intelligibility and improving overall quality of speech.

A. Basic Spectral Subtraction Algorithm

Spectral subtraction is one of the historic algorithm used for speech enhancement. It is simple and easy to implement, based on the principle that we can obtain an estimate of the clean signal spectrum by subtracting an estimate of the noise spectrum from the noisy speech spectrum. The noise spectrum can be calculated and altered, during the time interval when the signal is not there or when only noise is present. Assumption is noise is additive¸ its spectrum does not change with time means noise is stationary or it's slowly time varying signal¸ whose spectrum does not change significantly between the updating periods [8] [7]. When the speech signal is not present an estimate of the noise spectrum is subtracted from the noisy speech spectrum to obtain the estimated enhanced speech spectrum. The key advantage of this method of speech enhancement is that it is easy and easy to implement. Noise present in corrupted speech signal can be reduced effectively using spectral subtraction algorithm.

Let m(n) be the noisy speech signal given by[2]

$$p(n)=q(n)+r(n)$$

Where, q(n) represents the clean speech signal and r(n) is the uncorrelated additive noise. In spectral subtraction algorithm, to estimate the noise spectrum it is assumed that the noise and clean signal are uncorrelated. Initially, the spectral subtraction approach was used to estimate the short term magnitude spectrum of the clean signal by subtracting the estimated noise magnitude spectrum from the noisy signal magnitude spectrum . The noisy signal phase spectrum is used can estimate of the clean speech phase spectrum. The estimated time-domain clean

speech signal is obtained by taking the inverse Fourier Transform . However, this approach has several drawbacks. Therefore, another enhanced version of spectral subtraction algorithm is proposed, the averaging over multiple frames of a known noise segment. To recover the signal, the estimated magnitude spectrum is combined with the phase of the noisy signal and finally by applying Inverse Fourier Transform, the Clean speech is obtained .

Although the spectral subtraction algorithm can be easily implemented, yet, it has some drawbacks. The subtraction process needs to be done carefully to avoid any speech distortion. If too little is subtracted, much of the interfering noise remains, but if too much is subtracted, then some speech information might be lost[7].

B. Spectral Subtraction with over subtraction

To prevent the resultant spectrum from going below a preset minimum spectral floor value , an over estimate of the noise power is subtracted. This is the modification done to basic spectral subtraction and it leads to minimization of the perception of the narrow spectral peaks by reducing the spectral excursions and thus decreasing the impact of musical.

The SNR is defining as the ratio of noisy speech power to the estimated noise power. The parameter affects the amount of speech spectral distortion. If is too large, then more will be the attenuation of the stronger components with low signal to noise ratio (SNR). This prevents musical noise. If is too small then the noise remains in enhanced speech signal. Therefore we take the suitable value to prevent both musical and signal distortion. Parameter controls the amount of musical noise and residual noise. If is too small then musical noise will become audible but the residual noise will be reduced. If is too large then the residual noise will be audible but the musical noise related to spectral subtraction reduces.[6]

C. Non-linear spectral subtraction

Non linear subtraction algorithm [5] is proposed by Lockwood and Boudy. This is the modification done by making the over subtraction factor frequency dependent and the subtraction process non linear. In NSS algorithm we assume that the noise does not affects all spectral components equally. The low frequency region is more affected as compared to high frequency region by the certain type of noise due to this we suggests the use of frequency dependent subtraction factor for different types of noise. The process become non linear after using this factor subtraction. Larger values are subtracted at frequencies with low SNR levels and smaller values are subtracted at frequencies with high SNR levels.

Fig 1: Basic Flow Diagram for Speech Enhancement

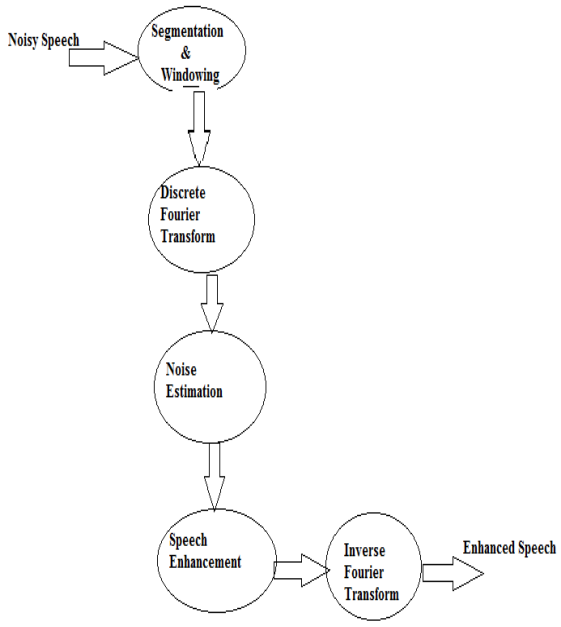

Table1: Comparison between different spectral subtraction

Basic Spectral Subtraction Method	Spectral Subtraction with over subtraction	Non Linear Spectral Subtraction (NSS)
An estimate of the clean signal spectrum is obtained by subtracting an estimate of the noise spectrum from the noisy speech spectrum.	In this method an over estimate is subtracted from the noise power spectrum and that prevents the resultant spectrum from going below a preset minimum level	In this method subtraction factor is made frequency dependent and subtraction process non-linear.
Assumption for this method is that noise is additive, its spectrum does not change with time means noise is stationary or it's slowly time varying signal, whose spectrum does not change significantly between the updating periods.	This method leads to minimizing the perception of the narrow spectral peaks by decreasing the spectral excursions and thus lower the musical noise effect.	Assumption for NSS is that noise does not affect all spectral component equally

III. LITERATURE REVIEW

Criteria	Past	Present
Name of Methods	Basic Spectral, Spectral Subtraction	Non-Linear Spectral Subtraction
Definition	These methods require subtracting of an estimated noise power spectrum from the speech power spectrum, by putting negative differences as zero, recombining the new power spectrum with the original phase, and then reconstructing the time waveform.[15]	This method uses the time trajectories of the short-time acoustic magnitude spectrum for the computation of the short-time modulation spectrum.[16]
Advantages	The spectral subtraction algorithm is computationally easy as it only includes a forward and inverse Fourier Transform. It also prevent the resultant spectral components from going below a preset minimum level (spectral floor)[15].	This method results in improved speech quality and does not introduce musical noise .
Disadvantages	While these method minimizes the broadband noise, it also introduces an annoying "musical noise"[3].	These methods involve large number of spectrum analysis.

Table 2: About Literature Review

Liuyang Gao et. al. [2] proposed that this paper presents and analyses a new speech enhancement algorithm based on improved spectral subtraction. Improved spectral subtraction algorithm accurately estimates the noise according to that the amplitude spectral of narrowband white Gaussian noise obeys Rayleigh distribution, based on that all noise can be changed into AWGN. This algorithm also adopts a new speech activity detection technology based on frequency band variance to detect signal activity. The emulational analyses indicates that the algorithm in this paper is better suit for speech enhancement by removing the noise in comparison to standard spectral subtraction.

Anuradha R. Fukane et. al. [3] describes that various spectral subtraction method in which the subtraction of noise spectrum from the noisy signal spectrum introduce a distortion in signal known as musical noise. The SNR of the noisy speech play an important role if SNR is less than 0dB no algorithm performed well. Female speech is less affected by noise. So the spectral subtraction algorithms are suitable for hearing aids in different noisy environments.

Anuprita P. Pawar et. Al. [4] proposes the review of various single channel speech enhancement methods in spectral domain. The authors say that the noise can have major impact on quality of the speech signal. If the noise is too low then the unwanted noise will be audible if it is too high then the speech will be distorted. It is observed that ESS method is suitable for noise reduction because it works in time domain and it is faster than frequency based method. The main advantage by using this NSS technique is that it does not require any voiced detection process by which performance of the system decreased.

Paurav Goel et. al. [1] discussed in this paper various techniques to reduce the musical noise from the noisy speech signal. This musical noise can be reduced to a certain limit by using the spectral subtraction techniques using modulation domain and geometric approach. When both subjective and objective test were performed on the modulation approach then we get the improved speech quality.

IV. CONCLUSION

This paper describes various spectral subtraction algorithms used for speech enhancement. These algorithms are simple and use FFT and Inverse FFT. Analysis of spectral subtractive algorithms revealed that these algorithms improve speech quality and not affect much more on intelligibility of speech signals. Non linear spectral subtraction (NSS) works in frequency domain. . So the spectral subtraction algorithms are suitable for reduction of noise in different environments with their modified versions. Different types of spectral subtraction algorithms were developed over the years. The most common variation considered the use of an over subtraction factor that controlled to some amount of speech spectral distortion caused by subtraction process.

REFERENCES

[1] Paruav Goel and Anil Garg “Review of spectral subtraction Technique for speech enhancement” IJECT Vol 2 December 2011.

[2] Liuyang Gao and Yunfei “Speech enhancement algorithm based on improved spectral suntraction “ IEEE 2009.

[3] Anuradha R. Fukane, Shashikant L. Sahare, "Enhancement of Noisy Speech Signals for Hearing Aids", 2011

[4] Anuprita P. Pawar , Kirtimalini B. Choudhari and Madhuri A Joshi , “Review of Single Channel Speech Enhancement Methods in Spectral Domain” International Journal of Applied Engineering Research, ISSN 0973-4562 Vol. 7 No.11 (2012)

[5]Lockwoord, P. and Boudy,J.,"Experiments with a Nonlinear Spectral Subtractor (NSS), Hidden Markov Models and the projection, for robust speech recognition in cars", Speech Communication, 11, pp215-228, Elsevier 1992.

[6] Ganga Prasad Surender, “A Review of Different Approaches of Spectral Subtraction Algorithms for Speech Enhancement”, Current Research in Engineering, Science and Technology (CREST) Journals,

[7]P.C. Loizou, “Speech Enhancement: Theory and Practice (CRC Press, Taylor Francis Group, Florida),” pp. 1-394, 2007.

[8]Yang Lu and Philipos C. Loizou, “A geometric approach to spectral subtraction”, ScienceDirect, Speech Communication,vol 50,pp.453-466, January 2008.

[9]Kim W, Kang S, and ko H.(2000) “Spectral subtraction Based on phonetic dependency and masking effects” IEEE Proc.vision image signal process, 147(5),pp423-427

[10]Virag, N., (1999). “Single channel speech enhancement based on masking properties of the human auditory system”.IEEE.Trans.Speech Audio Process,7(3), 126-137

[11]Ching-Ta Lu, “Enhancement of single channel speech using perceptual-decision-directed approach”, Science Direct, Speech Communication, vol.53, pp.495-507, December 2010.

[12]Yang Lu and Philipos C. Loizou, “A geometric approach to spectral subtraction”, ScienceDirect, Speech Communication,vol 50,pp.453-466, January 2008.

[13]Yi Hu and C. Loizou, “Evaluation of objective quality measure for speech enhancement”, IEEE trans. Audio and language processing, vol.6, no.1, January 2008.

[14]John H.L.Hansen and Bryan L. Pellom, “An objective quality evaluation protocol for speech enhancement algorithm”, The international conference on speech and language processing processing, 1998.

[15]Paruav Goel and Anil Garg ”Develpoments in Spectral Subtraction for Speech Enhancement”IJERA Vol 2 February 2012.

[16]Falk, T., Stadler, S., Kleijn, W.B., Chan, W.-Y., “Noise suppression based on extending a speech-dominated modulation band” In: Proc. ISCA Conf. Internat,2007

[17]Anuprita P Pawar, “Enhancement of Speech in Noisy Conditions” IJAREEIE July 2013.